



# FAILURE PROGNOSTICS BY A DATA-DRIVEN SIMILARITY-BASED APPROACH

Enrico Zio, F. Di Maio

## ► To cite this version:

Enrico Zio, F. Di Maio. FAILURE PROGNOSTICS BY A DATA-DRIVEN SIMILARITY-BASED APPROACH. International Journal of Reliability, Quality and Safety Engineering, 2013, 20, pp.1350001. 10.1142/S0218539313500010 . hal-00926377

**HAL Id: hal-00926377**

**<https://hal-centralesupelec.archives-ouvertes.fr/hal-00926377>**

Submitted on 9 Jan 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# FAILURE PROGNOSTICS BY A DATA-DRIVEN SIMILARITY-BASED APPROACH

Enrico Zio, Francesco Di Maio

*Energy Department, Polytechnic of Milan*

*Via Ponzio 34/3, 20133 Milano, Italy*

[enrico.zio@polimi.it](mailto:enrico.zio@polimi.it)

## ABSTRACT

This paper presents a data-driven, similarity-based approach for prognostics of industrial and structural components. The potentiality of the approach is demonstrated on a problem of crack propagation, taken from literature. The crack growth process is described by a non linear model affected by non-additive noises. A comparison is provided with a Monte Carlo-based estimation method, known as particle filtering.

### **Key Words:**

Prognostics, Residual Useful Life, Pointwise Fuzzy Similarity Analysis, Particle Filtering, Crack Growth.

# 1. Introduction

When the conditions of a component or structure can be monitored, maintenance can be planned dynamically on the basis of the observed conditions [Williams et al., 1994; Marseguerra et al., 2002]. This requires prognostic capability for predicting the evolution of the degradation state of the component or structure in the future. The aim of the prediction is to verify whether the component or structure can continue performing its function throughout its lifetime and, in case it cannot, to estimate the Remaining Useful Life (RUL), i.e. the time remaining before it can no longer perform its function [Jardine et al., 2006]. In practice, the estimate of the RUL of a component or structure may be difficult to obtain, since its degradation state may not be directly observable and/or the measurements may be affected by noise and disturbances.

Approaches to failure prognostics can be categorized broadly into model-based and data-driven [Chiang et al., 2001]. Model-based prognostics attempts to set up physical models of the component or structure for the estimation of the RUL. However, uncertainty due to the assumptions and simplifications of the adopted models may pose limitations on this approach.

Many researchers have focused on the problem of building exhaustive models of deteriorating components and structures to implement model-based prognostic tools. Markov and semi-Markov models have been widely exploited for achieving analytical results [Samanta et al., 1991; Lam et al., 1994; Hontelez et al., 1996; Kopnov, 1999; Yeh, 1997; Grall et al., 1998; Bérenguer et al., 2000]. On the basis of these models, several approaches have been proposed to analyze reliability-based and condition-based maintenance policies [Pulkkinen et al., 1992; Vlok et al., 2002; Castanier et al., 2002].

The most promising approaches rely on Bayesian methods to combine a prior distribution of the unknown degradation states with the likelihood of the observations collected, to build a posterior distribution [CUED-F-ENGTR310, 1998; Doucet et al., 2001]. In this setting, the estimation method most frequently used in practice is the Kalman filter, which is optimal for linear

state space models and independent, additive Gaussian noises [Anderson et al., 1979]. In this case, the posterior distributions are also Gaussian and computed exactly, without approximations. However, in most realistic cases the dynamics of degradation is non linear and/or the associated noises are non-Gaussian. Various approximate methods can be proposed to tackle these cases, e.g. the analytical approximations of the Extended Kalman (EKF) and the Gaussian-sum filters and the numerical approximations of the grid-based filters [Anderson et al., 1979]. Recently, Monte Carlo sampling methods are gaining popularity for their flexibility and ease of design [Kitagawa, 1987]. These methods go under the name of particle filtering because the continuous distributions of interest are approximated by a discrete set of weighed particles, where each particle represents a random trajectory of evolution in the state space and the weight is the probability of the trajectory [Djuric et al., 2003; Doucet et al., 2000; Cadini et al., 2009].

Data-driven techniques utilize monitored operational data related to system health. They can be beneficial when understanding of first principles of system operation is not straightforward or when the system is so complex that developing an accurate model is prohibitively expensive. Furthermore, recent advances in sensor technology and refined simulation capabilities enable us to continuously monitor the health of operating components and living structures, and to manage the related large amount of reference data.

Data-driven techniques can be divided into two categories: statistical techniques (regression methods, ARMA models, etc.) and Artificial Intelligence (AI) techniques (neural networks, fuzzy systems, etc.). The most direct data-driven techniques for RUL estimation attempts at fitting available data of component or structure degradation by regression models and then extrapolating the evolution up to failure. However, in practice, the component or structure degradation history available may be short and incomplete, and extrapolation may lead to large errors [Yan et al., 2004].

With respect to AI techniques, the most commonly used prediction methods are based on Neural Networks [Peel et al., 2008; Barlett et al., 1992; Santosh et al., 2009]. For prognostic tasks,

promising methods are Recurrent Neural Networks (RNNs) [Campolucci et al., 1999], Neuro-Fuzzy (NF) systems [Wang et al., 2004] and Support Vector Machines (SVMs) [Sotiris et al., 2007]. In spite of the recognized potential of empirical, data-driven techniques, limitations still exist for their use in safety critical applications, e.g. in nuclear technology, because of the lack of a systematic approach for selecting the structure and parameters of the models and their black-box character which limits intuition with respect to the understanding of their performance [Wang et al., 2008].

An opportunity for increased transparency and openness of empirical, data-driven models is offered by fuzzy logic-based methods, which are sustained by a formal mathematical framework for dealing with the vagueness of everyday reasoning [Zadeh, 1965] and accommodating measurement uncertainty and estimation imprecision [Yuan et al., 1997; Zio et al., 2005].

The subject of the present paper is the estimation of the RUL of a component or structure on the basis of measurements of its degradation state taken at predefined inspection times, which are possibly only few due to the fact that the lower the number of measurements, the lower the computational time and the cost associated to the inspection procedures. The computational framework proposed manipulates a set of degradation-to-failure patterns (hereafter called reference patterns) within a fuzzy-based, data-driven similarity analysis procedure [Angstenberger, 2001] for predicting the remaining life of a newly developing degradation pattern (hereafter called test pattern). The adoption of a fuzzy definition of pattern similarity allows capturing and integrating the possibly imprecise information carried by the measured degradation signal. An application is presented with reference to a non linear fatigue crack growth process, typical for a certain class of industrial and structural components [Oswald et al., 1984; Sobezyk et al., 1992; Bolotin et al., 1998; Myotyi et al., 2006].

The paper contents are structured as follows. Section 2 contains the description of the algorithm for fuzzy similarity analysis, at the basis of the prognostic approach to RUL estimation.

Section 3 presents the dynamic model of fatigue crack growth. In Section 4, the results of the application of the approach to the case study are presented and compared with those obtained by particle filtering. Finally, some conclusions on the advantages and limitations of the approach here propounded are given in Section 5.

## 2. Methodology

It is assumed that  $J$  measurements taken at predefined inspection times are available for  $N$  degradation-to-failure trajectories (reference patterns) of components or structures of the type of interest; these trajectories last all the way to component or structure failure, i.e., to the instance when the degradation state reaches the threshold value beyond which the component loses its functionality.

A degradation trajectory (test pattern) is developing in the actual component or structure under analysis, which is monitored at the predefined inspection times. The RUL estimation for the degrading component or structure is performed by analyzing the similarity between the test pattern and the  $N$  reference patterns, using their RULs weighted by how similar they are to the test pattern [Angstenberger, 2001]. The estimation procedure is adapted from a data-driven framework of recovery time prediction proposed by the authors for accident management [Zio et al., 2009].

Figure 1 shows a schematic sketch of the computational framework, with reference to degradation signal  $f(t)$ :

- **Step 1: fault detection.** The degradation signal  $f(t)$  is monitored throughout the time horizon of observation  $T$ , starting from (discrete) time  $t=1$ ; inspections of the component or structure degradation state, as indicated by signal  $f(t)$ , are made at predefined inspection times  $(T_1, T_2, T_3, \dots, T_J)$ , where  $T_j - T_{j-1} = n$  is the number of discrete time steps

time between two successive inspections; at each inspection time  $T_j$ , the value  $f(T_j)$  is recorded and appended to the vector of the values collected at the previous  $(j-1)$ -th inspections. For reasons which will become clear in the following, the reference patterns of signal values are organized in a reference matrix  $\overline{\overline{R}}_{[N \times n \cdot J]}$ , whose generic element  $r(i, k)$ ,  $k = 1, 2, \dots, n \cdot J$ , is the value assigned to the  $i$ -th reference pattern at time  $T_{j-1} < t_k < T_j$  by linear interpolation between the measured values  $r(i, n \cdot (j-1))$  and  $r(i, n \cdot j)$  at times  $T_{j-1}$  and  $T_j$ , respectively, and normalized in the range  $[0.2, 0.8]$ . The choice of the normalization range is motivated by the need to avoid numerical problems arising from manipulating null values.

As long as no signal trend is detected at inspection, the component or structure is qualified as working in nominal conditions and the estimate  $\hat{RUL}(T_j)$  of the remaining useful life made at the generic inspection time  $T_j$  is taken equal to the component or structure Mean Time to Failure  $MTTF(T_j)$ , computed from the remaining useful life  $RUL_i(T_j)$  of all the reference patterns:

$$\hat{RUL}(T_j) = MTTF(T_j) = \frac{1}{(|i| \parallel t_{f_i} > T_j)} \sum_{i \parallel t_{f_i} > T_j} (t_{f_i} - T_j) = \frac{1}{(|i| \parallel t_{f_i} > T_j)} \sum_{i \parallel t_{f_i} > T_j} RUL_i(T_j) \quad (1)$$

where  $t_{f_i}$  is the component or structure failure time of the  $i$ -th pattern (i.e., the time when the signal value exceeds the threshold beyond which the component or structure loses its functionality),  $(|i| \parallel t_{f_i} > T_j)$  is the cardinality of the set of reference patterns whose failure time is larger than  $T_j$  and  $RUL_i(T_j)$  is the remaining useful life on the reference pattern  $i$  starting from  $T_j$ . At the following inspection times, the algorithm continues to update the

estimate of  $\hat{RUL}(T_j) = MTTF(T_j)$  until a fault is detected upon inspection because the degradation signal has deviated from its nominal value beyond its range of allowed variability, which is a priori gauged with respect to the range of possible fluctuations in the signal; at this time, the RUL estimation algorithm is put to work to match the similarity of the developing degradation signal to the reference patterns and combine their failure times for estimating the component or structure RUL.

- **Step 2: pattern pointwise difference computation.** At the current time  $T_j$ , the latest  $n$ -long segment of values of the test pattern  $\bar{f}(T_j) = [f(T_j - n = T_{j-1}), f(T_j - n + 1), \dots, f(T_j - 1), f(T_j)]$  is built by linear interpolation of the values  $f(T_{j-1})$  and  $f(T_j)$  measured at the successive inspection times  $T_{j-1}$  and  $T_j$ . The test pattern  $\bar{f}$  is also normalized in  $[0.2, 0.8]$ . The pointwise difference  $\delta(\square)$  between the  $n$  normalized values of pattern  $\bar{f}$  and of the reference pattern segment  $r(i, k)$  is computed:

$$\delta(i, j) = \sum_{k=n \cdot (j-1)}^{n \cdot j} |\bar{f}(k) - r(i, k)|, \quad i = 1, 2, \dots, N, \quad j = 1, 2, \dots, J \quad (2)$$

The matrix  $\bar{\delta}_{[N \times J]}$  contains the difference measures  $\delta(i, j)$  between all  $n$ -long segments of the reference patterns and the test pattern.



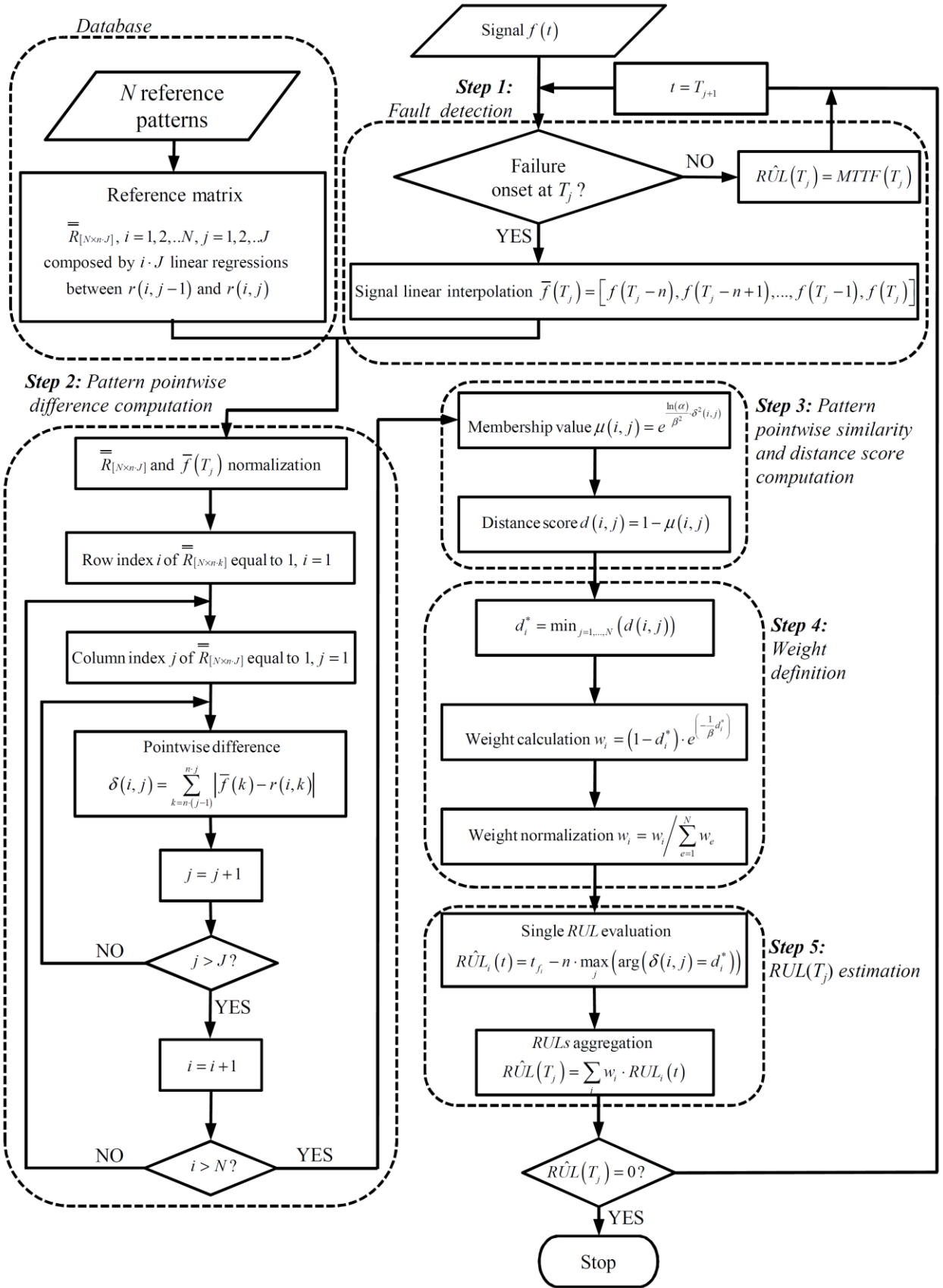


Figure 1 The flowchart of the fuzzy-based, data-driven approach

- **Step 3: pattern pointwise similarity and distance score computation.** In practice, there are numerous cases in which the similarity measure should allow for a gradual transition between ‘similar’ and ‘non-similar’ [Binaghi et al., 1993; Joentgen et al., 1999]. This can be achieved by resorting to a fuzzy logic modeling paradigm in which the pointwise difference of two patterns is judged for similarity with respect to an “approximately zero” fuzzy set (FS) specified by a function which maps the elements  $\delta(i, j)$  of the difference matrix  $\overline{\delta}_{[N \times J]}$  into their values  $\mu(i, j)$  of membership to the condition of “approximately zero”. The distance score  $d(i, j)$  between two pattern segments is then computed as:

$$d(i, j) = 1 - \mu(i, j), \quad i = 1, 2, \dots, N, \quad j = 1, 2, \dots, J \quad (3)$$

Common membership functions can be used for the definition of the FS, e.g. triangular, trapezoidal, and bell-shaped [Dubois et al., 1988]. In an exploration of the effects of these membership functions, all defined on the same support of the fuzzy set “approximately zero” and centered in the value zero, no particular sensitivity of the RUL estimation results has been recorded with respect to the particular shape. In the application illustrated in this work, the following bell-shaped function has turned out to give more robust results due to its gradual smoothness:

$$\mu(i, j) = e^{-\left(\frac{-\ln(\alpha)}{\beta^2} \delta^2(i, j)\right)} \quad (4)$$

The arbitrary parameters  $\alpha$  and  $\beta$  can be set by the analyst to shape the desired interpretation of similarity into the fuzzy set: the larger the value of the ratio  $\frac{-\ln(\alpha)}{\beta^2}$ , the narrower the fuzzy set and the stronger the definition of similarity. The choice of the values of  $\alpha$  and  $\beta$  depends on the application; one may proceed to determining the value  $\beta$  of the

difference value  $\delta$  which must have a degree of membership  $\mu$  equal to  $\alpha$  [Angstenberger, 2001; Zio et al., 2009].

- **Step 4: weight definition.** To assign the weight  $w_i$  given to the  $i$ -th reference pattern accounting for how similar it is to the test pattern, the minimum distance  $d_i^*$  along the  $i$ -th row of the matrix of Eq. (3) is first identified:

$$d_i^* = \min_{j=1,\dots,J} d(i, j), \quad i = 1, 2, \dots, N \quad (5)$$

The weight  $w_i$  is then computed, resorting to the arbitrarily chosen decreasing monotone function, which guarantees that the smaller the minimum distance the larger the weight given to the  $i$ -th reference pattern:

$$w_i = (1 - d_i^*) \cdot e^{\left(-\frac{1}{\beta} d_i^*\right)}, \quad i = 1, 2, \dots, N \quad (6)$$

The value of  $\beta$  in (7) is arbitrarily taken equal to that used in Eq. (4).

Then, the weight  $w_i$  is normalized:

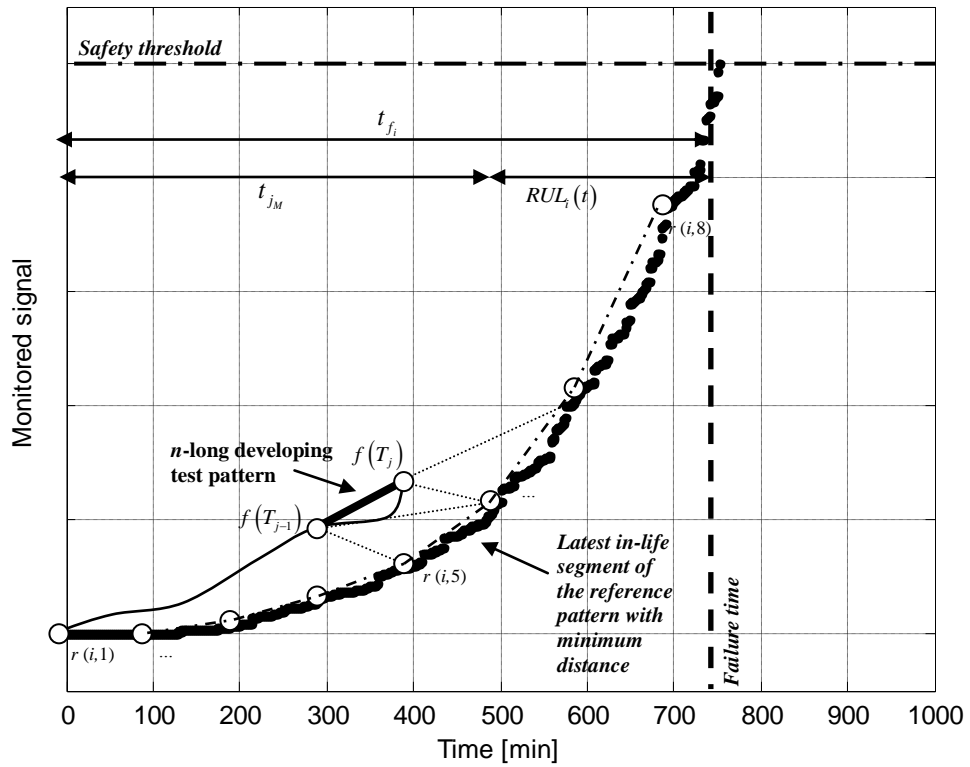
$$w_i = w_i / \sum_{e=1}^N w_e \quad (7)$$

- **Step 5:  $RUL_i(t)$  and  $RUL(T_j)$  estimation.** With respect to the generic  $i$ -th reference pattern for which  $t_{f_i} > t$ , the value  $RUL_i(t)$  is first determined as:

$$RUL_i(t) = t_{f_i} - t_{j_M}, \quad i = 1, 2, \dots, N \quad (8)$$

where  $t_{j_M} = n \cdot \max_j \left( \arg \left( \delta(i, j) = d_i^* \right) \right)$  is the final time of the latest-in-life segment of the  $i$ -th pattern among those which have minimum distance  $d_i^*$  from the developing test pattern ( $n$  is the test pattern length,  $\max_j \left( \arg \left( \delta(i, j) = d_i^* \right) \right)$  gives the largest column index  $j$  of  $\delta(i, \square)$

whose element is equal to  $d_i^*$ ). Thus,  $RUL_i(t)$  is the remaining time before failure on the reference pattern starting from the end time of the latest-in-life segment of minimum distance from the developing test pattern (Figure 2). This allows a conservative RUL estimation, biased towards “pessimistic” predictions of the RUL because in the case that more than one segment along the  $i$ -th reference pattern is closest to the developing test pattern, the latest one is taken, i.e. the one closest to failure.



**Figure 2** The remaining time before failure  $RUL_i(t)$  on a generic reference pattern starting from the end time of the latest in-life segment of minimum distance from the occurring test pattern

Then, at the generic inspection time  $T_j$ , the  $\hat{RUL}(T_j)$  is estimated as a similarity-weighted sum of the  $RUL_i(t)$ :

$$\hat{RUL}(T_j) = \sum_{i|t_{ji} > t} w_i \cdot RUL_i(t), \quad i = 1, 2, \dots, N \quad (9)$$

The ideas behind the weighting of the individual  $RUL_i(t)$  is that: *i*) all reference patterns bring useful information for determining the RUL of the degradation pattern currently developing; *ii*) those segments of the reference patterns which are most similar to the most recent segment of length  $n$  of the currently developing degradation pattern should be more informative in the extrapolation of the occurring pattern to failure.

### 3. The model of the fatigue crack growth process

Let us consider the process of crack growth in a component subject to fatigue. The common Paris-Erdogan model is adopted for describing the evolution of the crack depth  $x$  as a function of the load cycles  $t$  [Pulkkinen, 1991]:

$$\frac{dx}{dt} = C(\Delta S)^m \quad (10)$$

where  $C$  and  $m$  are constants related to the material properties [Provan, 1987; Kozin et al., 1989], which can be estimated from experimental data [Bigerelle et al., 1999] and  $\Delta S$  is the stress intensity amplitude, roughly proportional to the square root of  $x$  [Provan, 1987]:

$$\Delta S = \gamma\sqrt{x} \quad (11)$$

where  $\gamma$  is again a constant which may be determined from experimental data.

The intrinsic stochasticity of the process may be inserted in the model by modifying equation (10) as follows [Provan, 1987]:

$$\frac{dx}{dt} = e^{\omega} C (\gamma\sqrt{x})^m \quad (12)$$

where  $\omega \sim N(0, \sigma_\omega^2)$  is a white Gaussian noise. For  $\Delta t$  sufficiently small, the state-space model (12) can be discretized to give:

$$x(t_k) = x(t_{k-1}) + e^{\omega_j} C (\Delta S)^m \Delta t \quad (13)$$

which represents a non-linear Markov process with independent, non-stationary degradation increments of the degradation state  $x$ .

At the generic inspection time  $T_j$ , the degradation state  $x(T_j)$  is generally not directly measurable. In the case of non-destructive ultrasonic inspections a logit model for the observation  $f(T_j)$  can be introduced [Simola et al., 1998]:

$$\ln \frac{f(T_j)}{d - f(T_j)} = \gamma_0 + \gamma_1 \ln \frac{x(T_j)}{d - x(T_j)} + \nu_k \quad (14)$$

where  $d$  is the component material thickness,  $\gamma_0 \in (-\infty, \infty)$  and  $\gamma_1 > 0$  are parameters to be estimated from experimental data and  $\nu$  is a white Gaussian noise such that  $\nu \sim N(0, \sigma_\nu^2)$ .

The following standard transformations are introduced:

$$y(T_j) = \ln \frac{f(T_j)}{d - f(T_j)} \quad (15)$$

$$\mu(T_j) = \gamma_0 + \gamma_1 \ln \frac{x(T_j)}{d - x(T_j)} \quad (16)$$

In the case study here considered, taken from [Myotyrri et al., 2006], the parameters of the state equation (13) are  $C = 0.005$ ,  $m = 1.3$  and  $\gamma = 1$ , whereas those in the measurement Eq. (14) are  $\gamma_0 = 0.06$ , and  $\gamma_1 = 1.25$ . The process and measurement noise variances are  $\sigma_\omega^2 = 2.89$  and  $\sigma_\nu^2 = 0.22$ , respectively. The component is assumed failed when the crack depth  $x \geq d = 100$ , in arbitrary units. As an example, Figure 3 shows the degradation-to-failure pattern that in the following will be used as test pattern in the procedure for predicting the component RUL. The crack depth  $x$  reaches the full material thickness  $d = 100$  at 802 [min].

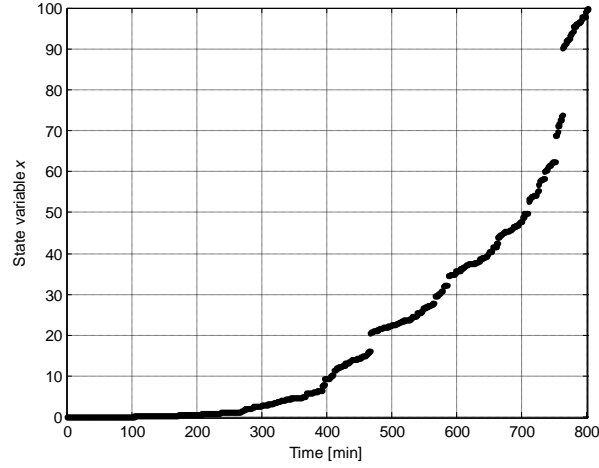


Figure 3 Crack growth pattern used as test pattern

## 4. Results

### 4.1 Application of the RUL estimation procedure on the fatigue crack propagation case study

The application of the procedure for computing at the inspection time  $T_j$  the estimate  $\hat{RUL}(T_j)$  (Section 2) on the fatigue crack propagation case study (Section 3) is hereafter illustrated. The interval between two successive inspections is equal to  $n_1 = 100$  [min] if the estimated  $RUL > 200$  [min], otherwise  $n_2 = 10$  [min], reflecting a more frequent inspection of the component or structure integrity as the component or structure is approaching the end of life. A database of  $N = 50$  reference crack propagation patterns of differing initial conditions is organized in the reference matrix  $\bar{\bar{R}}_{[N \times (n_1 + n_2) \cdot J]}$ , where  $N = 50$ ,  $J = 10$ ,  $n_1 = 100$  and  $n_2 = 10$ . The elements  $r(i, k)$  of such reference matrix are then compared for similarity with a test pattern containing the values of the measured signal of the developing degradation pattern. For each of the test patterns the procedural steps are performed as follows:

#### **Step 1: fault detection.**

The signal monitored starting from time  $t = 1$  [s] to the time horizon of observation  $T = 1000$  [min], with time step of  $n_1 = 100$  [min] until the estimated  $RUL > 200$  [min], or  $n_2 = 10$  [min] otherwise, is the observation  $f(t)$  of Eq. (14). This means that the actual safety parameter driving

the RUL, i.e., the state variable  $x$  of Eq. (13), is followed only indirectly through the observable variable, thus challenging further the robustness of the approach with respect to non-observable safety parameters. At each inspection time  $T_j$ ,  $j=1,2,...,30$ , its value is appended and stored in the vector containing the  $j-1$  values of the signal collected at the previous times. At the current time  $T_j$ , the latest  $n$ -long segment of values of the test pattern  $\bar{f}(T_j) = [f(T_j - n = T_{j-1}), f(T_j - n + 1), ..., f(T_j - 1), f(T_j)]$  is built by linear interpolation of the values  $f(T_{j-1})$  and  $f(T_j)$  measured at successive inspections. The Mean Time to Failure  $MTTF(T_j)$  is calculated resorting to Eq. (1) and  $\hat{RUL}(T_j)$  is set equal to  $MTTF(T_j)$  for each time step, until fault detection when the on-line data-driven, similarity-based RUL estimation is activated.

***Step 2: pattern pointwise difference computation.***

The pointwise difference  $\delta(i, j)$  between the test pattern  $\bar{f}(T_j) = [f(T_j - n = T_{j-1}), f(T_j - n + 1), ..., f(T_j - 1), f(T_j)]$  and  $r(i, k)$  is evaluated (Eq. (2)).

***Step 3: pattern pointwise similarity and distance score computation.***

The pointwise differences  $\delta(i, j)$  are mapped into values of membership  $\mu(i, j)$  of the “approximately zero” FS. The bell-shaped function of Eq. (4) is taken with parameters values  $\alpha = 0.2$  and  $\beta = 0.1$ . The distance scores  $d(i, j)$  are then computed by Eq. (3),  $i = 1, 2, ..., 50$ ,  $j = 1, 2, ..., 30$ .

***Step 4: weight definition.***

The minimum distances  $d_i^*$  are evaluated by Eq. (5), and the relative normalized weights  $w_i$  calculated through Eqs. (6) and (7),  $i = 1, 2, ..., 50$ .

***Step 5:  $RUL_i(t)$  and  $RUL(T_j)$  estimation.***



For each reference pattern, an estimate  $RUL_i(t)$  is computed by Eq. (8),  $i = 1, 2, \dots, 50$ ; then, the  $RUL_i(t)$  are aggregated in the weighted sum of Eq. (9) with the weights  $w_i$  previously calculated.

For the crack propagation pattern plotted in Figure 3, the estimates of the  $MTTF(T_j)$  are plotted in Figures 4 in thin continuous lines with the bars of one standard deviation of the samples  $(t_{f_i} - T_j | t_{f_i} > T_j)$ , where  $t_{f_i}$  is the time at which the crack depth  $x$  exceeds the material thickness  $d$  and thus failure occurs. The  $\hat{RUL}(T_j)$  estimates obtained based on pattern segments of  $n_1 = 100$  [min] and  $n_2 = 10$  [min], when  $\hat{RUL}(T_j) > 200$  [min] and  $\hat{RUL}(T_j) < 200$  [min], respectively, are plotted in bold circles; at the beginning of the test pattern, the predictions match the  $MTTF(T_j)$ ; then, upon fault detection, the  $\hat{RUL}(T_j)$  estimate moves away from the  $MTTF(T_j)$  values towards the real RUL (dashed thick line). In the Figure, the bold vertical line indicates the time of crack depth exceedance of the limit on the material thickness.

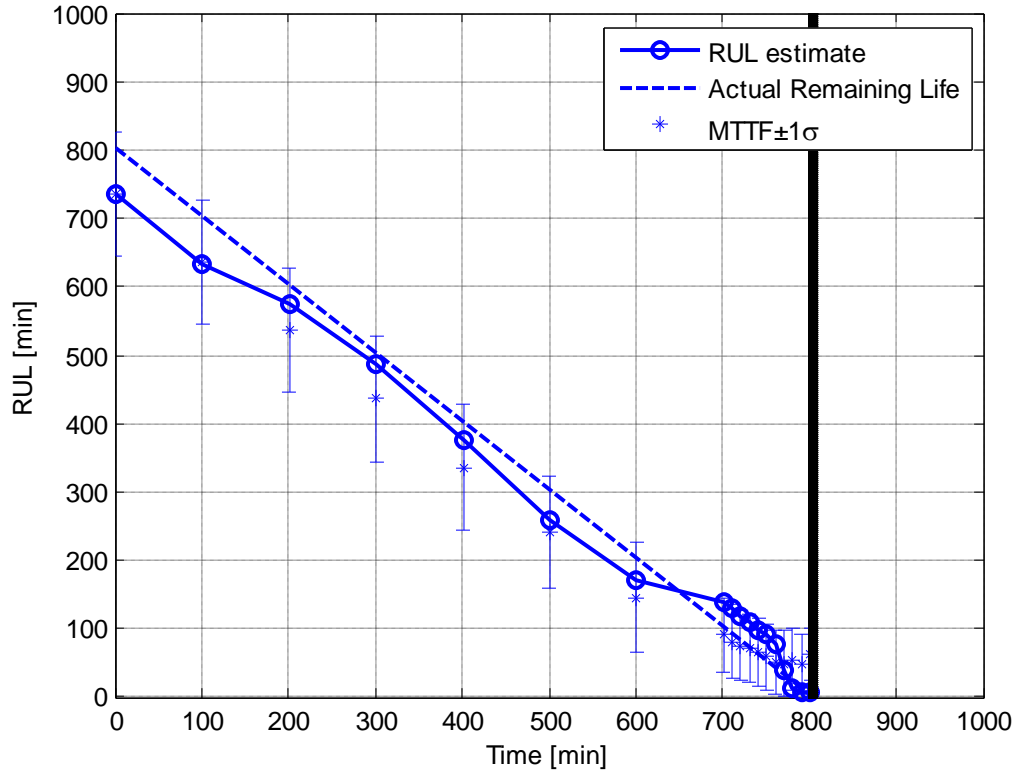


Figure 4 RUL estimation for the crack propagation pattern of Figure 3

#### 4.2 Comparison with a model-based particle filtering method

The Monte Carlo-based particle filtering method has been applied to the same case study of Section 3 in [Cadini et al., 2009]. The application of particle filtering for RUL estimation entails the evaluation of the conditional cumulative distribution function (cdf) of the stochastic observable variable related to the degradation state. For more details on the procedure, the interested reader may refer to [Cadini et al., 2009].

From Eq. (15) it follows that the transformed observation  $Y(T_j) \sim N(\mu(T_j), \sigma_v^2)$  is a Gaussian random variable with cdf:

$$cdf_{Y(T_j)}(y(T_j)|x(T_j)) = P(Y(T_j) < y(T_j)|x(T_j)) = \Phi\left(\frac{y(T_j) - \mu(T_j)}{\sigma_v}\right) \quad (17)$$

where  $\Phi(u)$  is the cdf of the standard normal distribution  $N(0,1)$ .

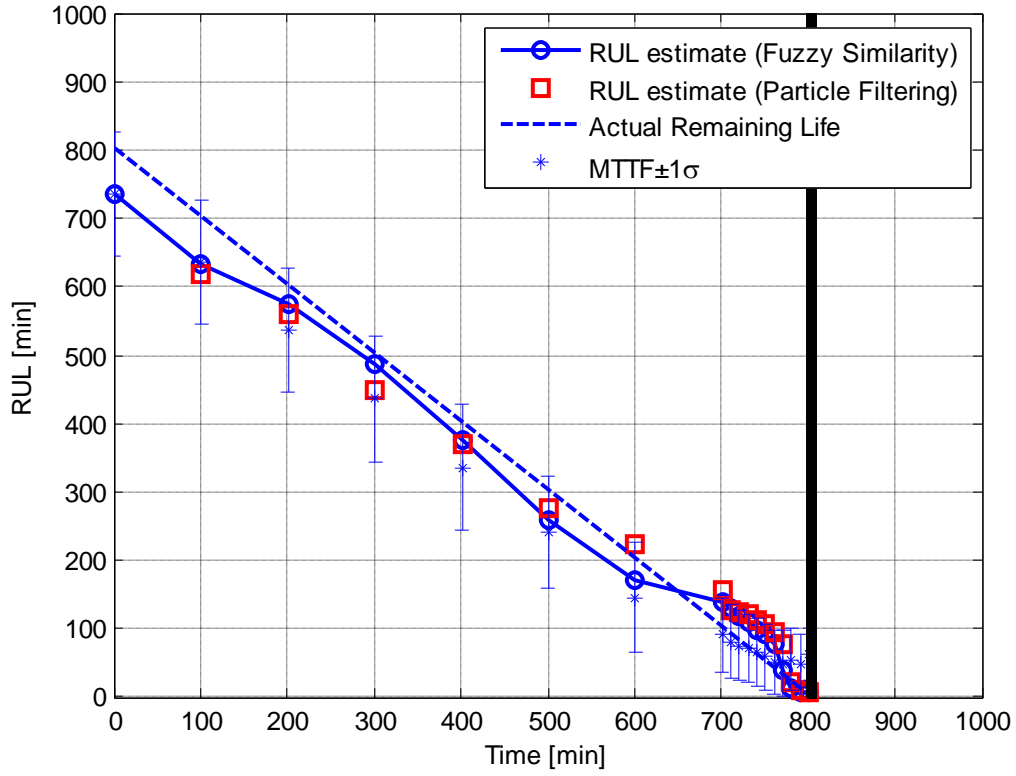
The conditional cdf of the stochastic measurement variable  $F(T_j)$  related to the stochastic degradation state  $X(T_j)$  is then:

$$cdf_{F(T_j)}(f(T_j)|x(T_j)) = cdf_{Y(T_j)}\left(\ln \frac{f(T_j)}{d-f(T_j)}|x(T_j)\right) = \Phi\left(\frac{1}{\sigma_v}\left(\ln \frac{f(T_j)}{d-f(T_j)} - \mu(T_j)\right)\right) \quad (18)$$

with corresponding probability density function (pdf):

$$pdf_{F(T_j)}(f(T_j)|x(T_j)) = \frac{1}{\sqrt{2\pi}\sigma_v} e^{-\frac{1}{2}\left(\frac{\ln \frac{f(T_j)}{d-f(T_j)} - \mu(T_j)}{\sigma_v}\right)^2} \frac{d}{f(T_j) \cdot (d-f(T_j))} \quad (19)$$

The estimates of the RUL obtained resorting to particle filtering are plotted in Figure 5 (together with those obtained by the data-drive, similarity-based approach of Section 2) in thin continuous lines with the bars of one standard deviation of the samples  $(t_{f_i} - T_j | t_{f_i} > T_j)$ ; the  $\hat{RUL}(T_j)$  estimates at the inspection times of Figure 4, are indicated in bold squares. The predictions provided by the two methods are comparable: after fault detection, the particle filtering  $\hat{RUL}(T_j)$  estimates similarly move away from the  $MTTF(T_j)$  values towards the real RUL (dashed thick line). In the Figure, the bold vertical line indicates the time of crack depth exceedance of the limit on the material thickness.



**Figure 5 Comparison of the RUL estimations for the crack propagation pattern of Figure 3 provided by the similarity based, data-driven approach and particle filtering**

## 5. Conclusions

A data-driven similarity-based prognostics procedure for estimating the RUL of a component or structure has been proposed. Data from different degradation-to-failure reference patterns are matched to the evolving degradation pattern based on a fuzzy pointwise similarity concept and their known residual life times are used for the estimation.

The prognostic procedure has been applied to the crack propagation dynamics of a component subject to fatigue cycles. Comparison with particle filtering demonstrates that the method is capable of successfully handling nonlinear dynamics and of dealing with non-Gaussian noises.

## References

- [Anderson et al., 1979] Anderson BD, Moore JB. Optimal filtering. Englewood Cliffs (NJ): Prentice Hall, 1979.
- [Angstenberger, 2001] Angstenberger, L., Dynamic Fuzzy Pattern Recognition, International Series in Intelligent Technologies, 17, Kluwer Academic Publishers, 2001.
- [Barlett et al., 1992] Barlett, E.B., Uhrig, R.E., Nuclear Power Plant Status Diagnostics Using an Artificial Neural Network, Nuclear Technology, 97, 1992.
- [Bérenguer et al., 2000] Bérenguer C, Grall A, Castanier B. Simulation and evaluation of conditionbased maintenance policies for multi-component continuous-state deteriorating systems. In: Cottam M, Harvey D, Pape R, Tait J, editors. Proceedings of the foresight and precaution conference, p. 275-82, 2000.
- [Bigerelle et al., 1999] Bigerelle, M., Iost, A., Bootstrap Analysis of FCGR, Application to the Paris Relationship and to Lifetime Prediction, International Journal of Fatigue, Vol. 21, pp. 299-307, 1999.
- [Binaghi et al., 1993] Binaghi, E., Della Ventura, A., Rampini, A., Schettini, R. (1993) Fuzzy Reasoning Approach to Similarity Evaluation in Image Analysis. International Journal of Intelligent Systems, Vol. 8, p. 749-769, 1993, 1993.
- [Bolotin et al., 1998] Bolotin VV, Babkin AA, Belousov IL. Probabilistic model of early fatigue crack growth. Prob Engng Mech, 13(3):227-32, 1998.
- [Cadini et al., 2009] Cadini, F., Zio, E., Avram, D., Monte Carlo-based filtering for fatigue crack growth estimation, Probabilistic Engineering Mechanics 24, 367-373, 2009.
- [Campolucci et al., 1999] Campolucci, P., Uncini, A., Piazza, F. and Rao, B.D., On-Line Learning Algorithms of Locally Recurrent Neural Networks, IEEE Trans. Neural Networks, 10, pp. 253-271, 1999.
- [Castanier et al., 2002] Castanier B, Bérenguer C, Grall A. Stochastic maintenance planning for a repairable system which is inoperative during maintenance operations. In: Bonano EJ, Camp AL, Majors MJ, Thompson RA, editors. Probabilistic safety assessment and management, PSAM6. Amsterdam: Elsevier, p. 1365-70, 2002.
- [Chiang et al., 2001] Chiang L. H., Russel, E., Braatz R., Fault detection and diagnosis in industrial systems, Springer-Verlag, London, 2001.
- [CUED-F-ENGTR310, 1998] Doucet A. On sequential simulation-based methods for Bayesian filtering, Technical report. Dept. of Engineering, University of Cambridge, CUED-F-ENGTR310, 1998.

- [Doucet et al., 2000] Doucet A, Godsill S, Andreu C. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statist Comput*, 10:197-208, 2000.
- [Doucet et al., 2001] Doucet A, de Freitas JFG, Gordon NJ. An introduction to sequential Monte Carlo methods. In: Doucet A, de Freitas JFG, Gordon NJ, editors. *Sequential Monte Carlo in practice*. New York: Springer-Verlag; 2001.
- [Djuric et al., 2003] Djuric PM, Kotecha JH, Zhang J, Huang Y, Ghirmai T, Bugallo MF, Miguez J. Particle filtering. *IEEE Signal Process*, 19-37, 2003.
- [Dubois et al., 1988] Dubois, D., Prade, H., Testemale, C., Weighted Fuzzy Pattern Matching. *Fuzzy Sets and Systems*, 28, 1988, p. 313-331, 1988.
- [Grall et al., 1998] Grall A, Bérenguer C, Chu C. Optimal dynamic inspection/replacement planning in condition-based maintenance. In: Lydersen S, Hansen G, Sandtorv H, editors. *Proceedings of the European safety and reliability conference ESREL'98*, p. 381-8, 1998
- [Hontelez et al., 1996] Hontelez JAM, Burger HH, Wijnmalen DJD. Optimum condition-based maintenance policies for deteriorating systems with partial information. *Reliab Engng Syst Safety*, 51:267-74, 1996.
- [Jardine et al., 2006] Jardine, A.K.S, Lin, D., Banjevic, D., A review on machinery diagnostics and prognostics implementing condition based maintenance, *Mech. Syst. and Sig. Process.*, Vol. 20, pp. 1483-1510, 2006.
- [Joentgen et al., 1999] Joentgen, A., Mikenina, L., Weber, R., Zimmermann, H.-J., Dynamic Fuzzy Data Analysis based on Similarity between Functions. *Fuzzy Sets and Systems*, 105 (1), 1999, p. 81-90, 1999.
- [Kitagawa, 1987] Kitagawa G. Non-Gaussian state-space modeling of nonstationary time series. *Journal of the American Statistical Association*, 82:1032-63, 1987.
- [Kopnov, 1999] Kopnov VA. Optimal degradation process control by two-level policies. *Reliab Engng Syst Safety*, 66:1-11, 1999.
- [Kozin et al., 1989] Kozin. F., Bogdanoff, J.L., Probabilistic Models of Fatigue Crack Growth: Results and Speculations, *Nuclear Engineering and Design*, Vol. 115, pp.143-71, 1989.
- [Lam et al., 1994] Lam C, Yeh R. Optimal maintenance policies for deteriorating systems under various maintenance strategies. *IEEE Trans Reliab*, 43:423-30, 1994.
- [Marseguerra et al., 2002] Marseguerra M, Zio E, Podofillini L. Condition-based maintenance optimization by means of genetic algorithms and Monte Carlo simulation. *Reliab. Eng. System Safety*, 77:151-65, 2002.
- [Myotyrä et al., 2006] Myotyrä E, Pulkkinen U, Simola K. Application of stochastic filtering for lifetime prediction. *Reliab. Eng. System Safety*, 91:200-8, 2006.

- [Oswald et al., 1984] Oswald GF, Schueller GI. Reliability of deteriorating structures. Eng Fracture Mech , 20(1):479-88, 1984.
- [Peel et al., 2008] Peel, L., Data Driven Prognostics using a Kalman Filter Ensemble of Neural Network Models, International Conference on Prognostics and Health Management, 2008.
- [Provan, 1987] J.W. Provan editor, Probabilistic fracture mechanics and reliability, Martinus Nijhoff Publishers, 1987.
- [Pulkkinen, 1991] Pulkkinen, U., A Stochastic Model for Wear Prediction through Condition Monitoring, K. Holmberg and A. Folkesson editors, Operational Reliability and Systematic Maintenance, London/New York: Elsevier, pp. 223-243, 1991.
- [Pulkkinen et al., 1992] Pulkkinen U, Uryas'ev S. Optimal operational strategies for an inspected component. In: Petersen KE, Rasmussen B, editors. Safety and reliability '92, proceedings of the European safety and reliability conference '92. London: Elsevier, p. 896-907, 1992.
- [Samanta et al., 1991] Samanta PK, Vesely WE, Hsu F, Subudly M. Degradation modeling with application to ageing and maintenance effectiveness evaluations. NUREG/CR-5612. US Nuclear Regulatory Commission, 1991
- [Santosh et al., 2009] Santosh, T.V., Srivastava, A., Sanyasi Rao, V.V.S., Gosh, A. K., Kushwaha, H.S., Diagnostic System for Identification of Accident Scenarios in Nuclear Power Plants using Artificial Neural Networks, Reliability Engineering and System Safety, 94, 759-762, 2009.
- [Simola et al., 1998] Simola, K, Pulkkinen, U., Models for non-destructive inspection data, Reliability Engineering and System Safety, Vol. 60, pp.1-12, 1998.
- [Sobezyk et al., 1992] Sobezyk K, Spencer BF. Random fatigue: From data to theory. Boston (MA): Academic Press; 1992.
- [Sotiris et al., 2007] V. Sotiris, M. Pecht, Support Vector Prognostics Analysis of Electronic Products and Systems, The AAAI Conference on Artificial Intelligence, 2007.
- [Wang et al., 2008] Wang T., Yu, J., Siegel, D., Lee, J., A similarity based prognostic approach for Remaining Useful life estimation of Engineered Systems, International Conference on Prognostics and Health Management, 2008.
- [Wang et al., 2004] Wang, W.Q., Goldnaraghi, M.F., Ismail, F., Prognosis of Machine Health Condition using Neuro-Fuzzy Systems, Mechanical Systems and Signal Processing, 18, 813-831, 2004.
- [Williams et al., 1994] Williams, J. H., Davies, A., Drake, I. R., Condition-based maintenance and machine diagnostics, Chapman and Hall, ISBN 0 4 12465 00 0, 187pp, 1994
- [Yan et al., 2004] Yan, J., Koç, M., Lee, J., A Prognostic Algorithm for Machine performance Assessment and its Application, Production Planning and Control, 15:8, 796-801, 2004.

- [Yeh, 1997] Yeh RH. State-age-dependent maintenance policies for deteriorating systems with Erlang sojourn time distributions. *Reliab Engng Syst Safety* ,58: 55-60, 1997.
- [Yuan et al., 1997] Yuan, B., Klir, G., Data driven identification of key variables, In: Ruan, D. (Ed.), *Intelligent Hybrid Systems Fuzzy Logic, Neural Network, and Genetic Algorithms*. Kluwer Academic Publishers, pp.161–187, 1997.
- [Vlok et al., 2002] Vlok PJ, Coetzee JL, Banjevic D, Jardine AKS, Makis V. An application of vibration monitoring in proportional hazards models for optimal component replacement decisions. *J OR Soc*,53:193-202, 2002.
- [Zadeh, 1965] Zadeh, L., Fuzzy sets, *Inform. Control* 8, 338–353, 1965.
- [Zio et al., 2005] Zio, E., Baraldi, P., Identification of nuclear transients via optimized fuzzy clustering, *Annals of Nuclear Energy*, 32, 1068–1080, 2005.
- [Zio et al., 2009] E. Zio, F. Di Maio, A Data-Driven Fuzzy Approach for Predicting the Remaining Useful Life in Dynamic Failure Scenarios of a Nuclear Power Plant, *Reliability Engineering and System Safety*, RESS, 10.1016/j.ress.2009.08.001, 2009.